A Survey Paper on Muti-Hop Privacy-Aware Data Aggregation in Mobile Sensing

D. N. Rewadkar[1], Asmita D. Abhyankar[2]

[1]H.O.D, Department of Computer Engineering, RMD Sinhgad School of Engineering, University of Pune, India

[2]Department of Computer Engineering, RMD Sinhgad School of Engineering, University of Pune, India

Abstract: ***With more increasing capabilities of mobile devices or we can say smart phones give rise to a variety of mobile sensing applications. This paper specifies how an untrusted aggregator in mobile sensing can periodically gain desired statistics over the data contributed by multiple mobile users, with privacy of each user. Although there are some existing works in this area, they either require both directional communications between the aggregator and mobile users in every aggregation period, or have high-computation overhead and cannot support large plaintext spaces. Also, they do not consider the Min aggregate, which is quite useful in mobile sensing. To address these problems, we propose an efficient protocol to obtain the Sum aggregate, which employs an additive homomorphic encryption and a novel key management technique to support large plaintext space. We also extend the sum aggregation protocol to obtain the Min aggregate of time-series data. To deal with dynamic joins and leaves of mobile users, a scheme proposed that utilizes the redundancy in security to reduce the communication cost for each join and leave***

Keywords: mobile sensing, privacy, data aggregation

1. Introduction

Mobile devices such as smart phones are gaining an ever-increasing popularity. Most smart phones are equipped with a rich set of embedded sensors such as camera, microphone, GPS, accelerometer, ambient light sensor, gyroscope, and so on. The data generated by these sensors provide opportunities to make sophisticated inference about not only people (e.g., human activity, health, location, social event) but also their surrounding (e.g.,pollution, noise, weather, oxygen level), and thus can help improve people's health as well as life. This enables various mobile sensing applications such as environmental monitoring, traffic monitoring, healthcare, and so on.

In many scenarios, aggregation statistics need to be periodically computed from a stream of data contributed by mobile users, to identify some phenomena or track some important patterns. For example, the average amount of daily exercise (which can be measured by motion sensors) that people do can be used to infer public health conditions. The average or maximum level of air
pollution and pollen concentration in an area may be useful for people to plan their outdoor activities. Other statistics of interests include the lowest gasoline price in a city, the highest moving speed of road traffic during rush hour, and so on.

In mobile Network it is sometimes necessary for users to share the power to use a cryptosystem. The system secret is divided up into shares and securely stored by the entities forming the distributed cryptosystem. The main advantage of a distributed cryptosystem is that the secret is never computed, reconstructed, or stored in a single location, making the secret more difficult to compromise.

Investigations within the fields of threshold group-oriented aggregated Key schemes, threshold group aggregated Key schemes, Multisink Time Stamp schemes, and Threshold-Multisink Time Stamp schemes resulted in explicitly defining the properties of Threshold-Multisink Time Stamp schemes.

2. Related Work

Set Many works have addressed various security and privacy issues in mobile sensing networks and systems but they do not consider data aggregation. There are a lot of existing works on security and privacy-preserving data aggregation,but most of them assume a trusted aggregator and cannot protect user privacy against untrusted aggregators. Yang et al. proposed an encryption scheme that allows an untrusted aggregator to obtain the sum of multiple users's data without knowing any specific user's data. However, their scheme requires expensive rekeying operations to support multiple time steps, and thus may not work for time-series data.

Shi et al. proposed a privacy-preserving data aggregation scheme based on data slicing and mixing techniques. However, their scheme is not designed for time-series data. It may not work well for time-series data,since each user may need to select a new set of peers in each aggregation interval due to mobility. Besides, their scheme for nonadditive aggregates (e.g., Max/Min) requires multiple rounds of bidirectional communications between the aggregator and mobile users which means long delays. In contrast, our scheme obtains those aggregates with just one round of unidirectional communication from users to the aggregator.

Rieffel et al. proposed a construction that does not require an extra round of interaction between the aggregator and the users. In their scheme, the computation and storage cost is roughly equal to the number of colluding users that the system can tolerate. Thus, their scheme has high overhead to achieve good resistance to collusion, especially when the system is large and a large number of users collude. In contrast, our scheme tolerates a high fraction of colluding users (e.g., 30 percent) with very small cost even when the system is large. Acs and Castelluccia also proposed a scheme based on additive homomorphic encryption, but in their scheme each node shares a pairwise key with any other node. Shi et al. proposed a construction for sum aggregation based on the assumption that the Decisional Diffie-Hellman problem is hard over finite cyclic groups. In their construction, each user sends her ciphertext to the aggregator and no communication is needed from the aggregator to the users.

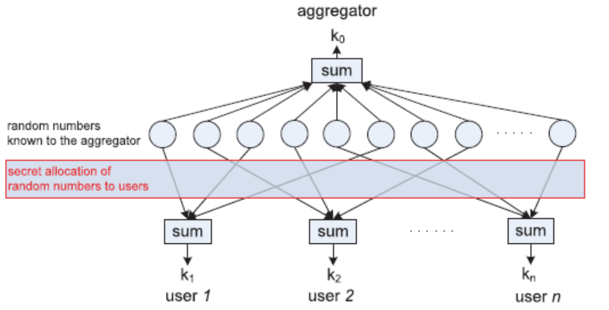
To decrypt the sum, their construction needs to traverse the possible plaintext space of sum, and thus, it is not efficient for a large system with large plaintext spaces. Chan et al. extended the construction in with a binary interval tree technique, but their scheme still has the limitation in plaintext spaces. Jawurek and Kerschbaum proposed a scheme that provides differential privacy for sum. Our aggregation protocol for sum can be used as a building block of their scheme to improve the computational efficiency.

3. Proposed System

This project is to propose a new Multisink Time Stamp scheme without a trusted third party (TTP), based on a round optimal, publicly verifiable DKG protocol. The proposed scheme can be easily adapted to incorporate a TTP; a version of the proposed scheme with the assistance of a TTP will therefore not be presented.

The proposed discrete logarithm-based Multisink Time Stamp scheme is also proactively secure, allowing for DKR to a new access structure and periodic DKU to mitigate attacks from an active/mobile adversary. The proposed discrete logarithm-based Multisink Time Stamp scheme is made proactively secure by periodically updating secret shares and facilitating changes in group membership by allowing an authorized subset of existing group members to redistribute secret shares to a new access structure.

The scheme fulfills all the fundamental properties of generic Multisink Time Stamp schemes given in the properties of Multisink Time Stamp and resists attacks to which other similar schemes are subject. The straw-man construction is as follows:



The intuition behind the straw-man construction. The aggregator computes the sum of a set of random numbers as the decryption key. These numbers are secretly allocated to the users, and each user computes the sum of its allocated numbers as the encryption key. The aggregator does not know which random numbers are allocated to each user, and thus does not know any user's key.

3.1 Intuition

Intuition of the straw-man construction. Suppose there are nc random numbers. The aggregator has access to all the numbers, and it computes the sum of these numbers as the decryption key k0. These numbers are divided into n random disjoint subsets, each of size c. These n subsets are assigned to the n users, where each user has access to one subset of numbers. User i compute the sum of the numbers assigned to it as the encryption key ki. Clearly, holds. The aggregator cannot know any user's encryption key because it does not know the mapping between the random numbers and the users. When c is large enough, it is infeasible for the aggregator to guess the numbers assigned to a particular user with a brute-force method. The aggregator's decryption key cannot be revealed by any user because no user knows all the numbers.

3.2 Construction

The construction is as follows: Secret Setup. The key dealer generates nc random and different secrets s1snc. It divides these secrets into n random disjoint subsets, with c secrets in each subset. Let S denote the set of all secrets,

3.3 Group Public Key Length

The Multisink Time Stamp scheme avoids conspiracy attacks without attaching a random secret to shares. The group public key is dependent on the number of group members, as the aggregated Key verifier needs the individual public values of all group members to compute the subgroup public key that is required to verifying the aggregated Key. Difficulty will be experienced with this scheme when trying to eliminate the need for a trusted authority to distribute the initial group key shares.

A robust authentication mechanism is essential for securing a distributed system against active adversaries and central to ensure the traceability of individual Signers. The proposed Multisink Time Stamp scheme uses the long-term private keys of the members, provided by a public key infrastructure, to avoid conspiracy attacks even if colluding members derive or control the group secret. As a result of members including their private keys in their individual aggregated Keys, the public key of the scheme consists of the public key of the subgroup that collaborated to generate the threshold aggregated Key. The public key of the subgroup is a function of the long-term public keys of the group members.

Although the group public key may be perceived to be dependent on the group size, the scheme does not introduce any additional storage requirements since the public keys used in the calculation is publicly known (traceable) and primarily required for authentication purposes.

3.4 Group-Oriented Aggregated Key Size

The main contribution to the communication overhead, post aggregated Key generation, is made by the size of the group aggregated Key. The aggregated Key size of Multisink Time Stamp schemes is bound to be dependent on the threshold parameter. This conclusion is drawn from the traceability property of Multisink Time Stamp schemes, which specifies that any outsider must be able to retrieve the identities of the individual signers from the threshold aggregated Key.

The threshold aggregated Key must thus be bound to information explicitly linked to each of the signers that collaborated to generate the threshold aggregated Key. In the case of the proposed scheme, the information is the identities of the individual signers. The individual identities of the group members can be carefully chosen to significantly reduce the size of the Multisink Time Stamp.

3.5 Communication Cost of aggregated key Generation and Verification

In terms of communication cost, the individual and threshold aggregated Key generation mechanisms of all the existing Multisink Time Stamp schemes and the proposed scheme are almost equivalent. Multiparty aggregated Key schemes constructed from Straw-man type (discrete logarithm-based) aggregated Key variants are bound to be interactive.

In round one, each participant generates a commitment and in the second round, generates an individual aggregated Key on an arbitrary message. In the third round, participants send their contribution to a combiner or designated clerk which constructs the threshold aggregated Key.

Assume the authorized subset of group members collaborate to sign a message. This yields a three round protocol for existing schemes, which requires broadcast messages and unicast messages. The proposed Multisink Time Stamp scheme, is to the best of all other schemes. The proposed scheme also eliminates the need for a combiner. Assume that the group contains at least one malicious or faulty participant, the proposed protocol will still require three rounds and only two rounds if all individual aggregated Keys are verified.

3.6 Computational Cost of Aggregated Key Generation and Verification

To make a feasible comparison between the computational cost of the proposed Multisink Time Stamp scheme and similar schemes it is assumed that the system parameters are chosen to yield the same time complexity for exponentiations, multiplications, and summations. Although summations and, in some cases, multiplications contribute to an insignificant fraction of the overall time complexity, these operations are still included for the sake of completeness.

Values that remain constant between different aggregated Key generations can be precomputed and are therefore not included in the analysis. The computational cost of the schemes will be given in terms of the minimum members required to Collaboratively sign an arbitrary message.The computational overhead that causes the most concern is the number of exponentiations in the individual aggregated Key verification and in Multisink Time Stamp verification, which are anticipated to contribute the bulk of the verification time complexity.

The justification for looking critically at the verification processes is substantiated by the notion that a aggregated Key is normally generated only once, but verified many times. The optimum number of exponentiations for an Straw-man type aggregated Key variant is 2.It can thus be concluded that the proposed Multisink Time Stamp scheme is superior to existing schemes since it requires only two exponentiations for Multisink Time Stamp verification, while guaranteeing break-resistance. For individual aggregated Key verification, three exponentiations are required, one more than the optimal two exponentiations. The additional exponentiation is as a consequence of satisfying the stronger break-resistance property.

4. Conclusion

The main aim of this system is to introduce a secure Multisink Time Stamp scheme. To reach this objective, the secure and optimally efficient Straw-man type aggregated Key variant, GES, was extended to a multiparty setting to yield a Multisink Time Stamp scheme, which provides a guaranteed traceability property. The proposed Multisink Time Stamp scheme was shown to satisfy all of the specified security requirements and fulfills the stronger break-resistant property. The Multisink Time Stamp aggregated Key scheme thus remains secure, even if the threshold cryptosystem has been broken, i.e., the group secret or individual secret shares are known or controlled by an adversary.

The efficiency analysis showed that the proposed Multisink Time Stamp scheme outperforms other existing schemes and is optimal in terms of exponentiations with respect to threshold aggregated Key verification and near optimal for individual aggregated Key verification, while providing break resistance.

Use of the DKRU mechanism makes the proposed fully distributed Multisink Time Stamp scheme proactively secure, allows for dynamic group membership, and gives the group members the capability of adjusting the security trade-off by redistributing the existing access structure to a new access structure

References

[1] Qinghua Li,Guohong Cao,Thomas F. La Porta, “Efficient and Privacy-Aware Data Aggregation in Mobile Sensing”,IEEE Transactions on dependable and secure computing.

[2] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M.Hansen, E. Howard, R. West, and P. Boda, “Peir, the Personal Environmental Impact Report, As a Platform for Participatory Sensing Systems Research,” Proc. ACM/USENIX Int’l Conf. Mobile Systems, Applications, and Services (MobiSys ’09), pp. 55-68, 2009.

[3] J. Hicks, N. Ramanathan, D. Kim, M. Monibi, J. Selsky, M. Hansen, and D. Estrin, “AndWellness: An Open Mobile System for activity and Experience Sampling,” Proc. Wireless Health, pp. 34-43, 2010.

[4] N.D. Lane, M. Mohammod, M. Lin, X. Yang, H. Lu, S. Ali, A.Doryab, E. Berke, T. Choudhury, and A. Campbell, “Bewell: ASmartphone Application to Monitor, Model and Promote Wellbeing,” Proc. Fifth Int’l ICST Conf. Pervasive Computing Technologiesfor Healthcare, 2011

[5] V. Rastogi and S. Nath, “Differentially Private Aggregation of Distributed Time-Series with Transformation and Encryption,” Proc. ACM SIGMOD Int’l Conf. Management of Data, 2010.

Author Profile

Prof. D. N. Rewadkar received M.E. Computer Technology, from S.R.T.M. University, Nanded (2000).Currently he is working as the H.O.D of Computer Engineering Department in RMD SSOE,Warje, Pune. He was a Member of Board of Study committee of S.R.T Marathwada University, Nanded for Computer Science & Engineering. He has 21 years of teaching experience.

Asmita D. Abhyankar Research Scholar RMD Sinhgad School of Engineering, University of Pune. She received B.E. in Information Technology from Information Technology department of Pune Vidyarthi Griha’s College of engineering and technology from University of Pune, Pune). Currently she is pursuing M.E. in Computer Engineering from RMD Sinhgad School of Engineering, Warje, Pune, University of Pune.